AD-A087 117 IOWA UNIV IOWA CITY DEPT OF STATISTICS F/G 12/1
BOUNDS ON MIXTURES OF DISTRIBUTIONS ARISING IN ORDER RESTRICTED--ETC(U)
JUL 80 T ROBERTSON, F T WRIGHT N00014-80-C-0322
UNCLASSIFIED 1 NL

1 OF 1
AD A
087117

END
DATE
FILMED
9-80
DTIC

LEVEL

BOUNDS ON MIXTURES OF DISTRIBUTIONS ARISING IN ORDER RESTRICTED INFERENCE[1]

(12)

Tim Robertson and F. T. Wright
University of Iowa, University of Missouri - Rolla

Abbreviated Title: Distributional Bounds

American Mathematical Society 1970 subject classification.
Primary 62E15 Secondary 62G10

Key words and phrases:

Order restricted inference, tests for and against a trend, chi-bar-squared distribution, E-bar-squared distribution , tail probability bounds, least favorable configurations.

(1)This research was partially sponsored by the Office of Naval Research under Contracts ONR N00014-80-C0322 and ONR N00014-80-C-0321.

ABSTRACT

In testing hypotheses involving order restrictions on a collection of parameters, distributions arise which are mixtures of standard distributions. Since tractable expressions for the mixing proportions do not, in general, exist even for parameter collections of moderate size, the implementation of these tests may be difficult. It is not necessary to bound the individual mixing proportions in order to obtain a bound for the tail probabilities of the "mixed" distributions. Bounds for such tail probabilities are derived and their application to several order restricted hypothesis testing situations is discussed. These results are also applied to obtain the least favorable configuration in a two sample likelihood ratio test of the equality of two multinomial populations versus a stochastic ordering alternative.

1. INTRODUCTION. Distributions having tail probabilities, which are weighted averages of the tail probabilities of standard distributions, play an intrinsic role in order restricted hypothesis tests (cf. Barlow et al. (1972), particularly Chapter 3). Use of these test procedures can be cumbersome since no tractable expressions exist, in general, for the weights associated with the various standard tail probabilities. In order to be more specific, consider the following example. Suppose $\bar{X}_1, \bar{X}_2, \ldots, \bar{X}_k$ are the means of independent random samples from normal distributions with means $\mu_1, \mu_2, \ldots, \mu_k$ and variances $\sigma_1^2, \sigma_2^2, \ldots, \sigma_k^2$, respectively. If the variances are known, then a likelihood ratio statistic for testing H_0: $\mu_1 = \mu_2 = \ldots = \mu_k$ vs. H_1: $\mu_1 \leq \mu_2 \leq \ldots \leq \mu_k$ is

$$T_{01} = -2 \ln \lambda = \sum_{i=1}^{k} w_i(\bar{\mu}_{ki} - \hat{\mu})^2$$

where λ is the likelihood ratio, $w_i = n_i/\sigma_i^2$, n_i is the size of the ith sample, $\hat{\mu} = \sum_{i=1}^{k} w_i \bar{X}_i / \sum_{i=1}^{k} w_i$ and $\bar{\mu}_{k1}, \bar{\mu}_{k2}, \ldots, \bar{\mu}_{kk}$ are the maximum likelihood estimators of $\mu_1, \mu_2, \ldots, \mu_k$ subject to the restriction imposed by H_1. (Computation of the $\bar{\mu}_{ki}$ is discussed at length in Barlow et al. (1972).) Bartholomew (1959) studied T_{01} and proved that

$$(1) \qquad P[T_{01} \geq c] = \sum_{\ell=1}^{k} P(\ell,k)\, P[\chi^2_{\ell-1} \geq c]$$

whenever H_0 is satisfied. The symbol, χ^2_ν, will be used to denote a chi-squared variable with ν degrees freedom ($\chi_0^2 \equiv 0$) and $P(\ell,k)$ denotes the probability, under H_0, that there are

exactly ℓ distinct values (levels) among $\bar{\mu}_{k1}, \bar{\mu}_{k2}, \ldots, \bar{\mu}_{kk}$. The distribution given in (1) is called a chi-bar-squared distribution.

A detailed discussion of the $P(\ell,k)$s may be found in Barlow et al. (1972). (All references in this and the next two paragraphs refer to that source.) Use of T_{01}, as a test statistics, is cumbersome, for unequal weights and for $k \geq 5$, since no tractable general expression for $P(\ell,k)$ has been found. Explicit formulas for $k \leq 4$ are given on pages 140 and 141. If the weights are all equal, the $P(\ell,k)$ can be found recursively from Corollary B on page 145 and their values in this important special case are tabled in Appendix A.5 for $k \leq 12$. In theory, the recursion formula, (3.23), on page 139 gives a general tool for computing the $P(\ell,k)$s. However, its implementation is virtually impossible since it requires computation of $P(j,j)$ and for $j \geq 5$ no closed form expression for $P(j,j)$ exists. (By interpolating in the table given in Abrahamson (1964), one can obtain $P(5,5)$.) Of course, for fixed k and specific weights one could Monte Carlo the $P(\ell,k)$.

It has been suggested that the $P(\ell,k)$s are fairly robust to the weights (cf. Grove (1980)) and that the values for equal weights give reasonable approximations except in unusual cases. However, the nature of this robustness does not seem to have been quantified and no insight into the nature of those rare cases has been given. Insight into both of these claims is given by the analysis in Section 2.

The bound $P(k,k) \leq (1/2)^{k-1}$ is given on page 138. Similar

bounds for $P(\ell,k); \ell = 1,2,\ldots,k-1$ can be obtained as follows. For arbitrary positive weights, using the minimum lower sets algorithm for $\bar{\mu}_{k1},\bar{\mu}_{k2},\ldots,\bar{\mu}_{kk}$, we have

$$(2) \qquad P(1,k) = P\left[\min_{1 \le \alpha \le k} \frac{\sum_{j=1}^{\alpha} w_j \bar{X}_j}{\sum_{j=1}^{\alpha} w_j} = \frac{\sum_{j=1}^{k} w_j \bar{X}_j}{\sum_{j=1}^{k} w_j}\right]$$

$$\le P\left[\frac{\sum_{j=1}^{k-1} w_j \bar{X}_j}{\sum_{j=1}^{k-1} w_j} \ge \frac{\sum_{j=1}^{k} w_j \bar{X}_j}{\sum_{j=1}^{k} w_j}\right] = 1/2.$$

These bounds for $P(j,j)$ and $P(1,j)$ can be used in the recursion formula, (3.23), to derive the following:

$$(3) \qquad P(\ell,k) \le \sum_{r=1}^{k-\ell} \binom{\ell}{r}\binom{k-\ell-1}{r-1}(1/2)^{\ell+r-1}; \; 1 \le \ell \le k-1.$$

This bound seems fairly reasonable for values of ℓ close to k but not so sharp for smaller values of ℓ. For example, it yields $P(k-1,k) \le (k-1)/2^{k-1}$ which seems as sharp as the bound for $P(k,k)$. On the other hand, for $k = 5$ it yields 8/16, 12/16, 9/16, as bounds for $P(1,5)$, $P(2,5)$, and $P(3,5)$ respectively.

As we shall see in Section 2, we do not need to bound the individual $P(\ell,k)$s in order to get a bound for (1). Specifically, in order to obtain an upper bound for (1) it sufficies to find numbers $b_1,b_2,\ldots,b_k$ such that $\sum_{\ell=\alpha}^{k} b_\ell \ge \sum_{\ell=\alpha}^{k} P(\ell,k)$ for $\alpha = 1,2,\ldots,k$. Bounds of this form are given in Theorem 1 and Remarks 2 and 3 show that these bounds cannot be improved, in general. In Section Three application of these bounds to several order restricted testing problems is discussed. One application

considered is to a likelihood ratio test of the equality of two multinomial populations against a stochastic ordering alternative. Here, the bounds give a least favorable configuration within the non-simple null hypothesis and can be used to describe conservative rejection regions for this test. Details of this problem are described in another paper.

The theory in Section 2 gives one a feel for the sensitivity of the $P(\ell,k)$ to the weight, as well as the circumstances under which one could either use the bounds or the expressions for the equal weights case as approximations.

2. TAIL PROBABILITY BOUNDS. The vector $\bar{\mu}_k = (\bar{\mu}_{k1},\bar{\mu}_{k2},\ldots,\bar{\mu}_{kk})$ is referred to as the isotonic regression of the vector $\bar{X} = (\bar{X}_1,\bar{X}_2,\ldots,\bar{X}_k)$ with weights $w_1,w_2,\ldots,w_k$. The computation algorithms referred to in this section are described in detail in Barlow et al. (1972). A basic idea in this section is that since $P[\chi^2_\nu \geq c] \leq P[\chi^2_{\nu+1} \geq c]$ for $\nu = 0,1,2,\ldots$ and for all c, one obtains an upper (lower) bound for (1) by making the distribution of the number of level sets in $\bar{\mu}_k$ as large (small) as possible.

Theorem 1. Assume that $\mu_1 = \mu_2 = \ldots = \mu_k = \mu$. If $a_1 \leq a_2 \leq \ldots \leq a_k$, then

$$(4) \qquad (a_1+a_2)/2 \leq \sum_{\ell=1}^{k} P(\ell,k)a_\ell \leq \sum_{\ell=1}^{k} \binom{k-1}{\ell-1} 2^{-k+1} a_\ell$$

for $k = 2,3,\ldots$. If $a_1 \geq a_2 \geq \ldots \geq a_k$, then the inequalities in (4) are reversed.

Proof: The second conclusion follows from the first by noting

that, in this case, $-a_1 \le -a_2 \le \dots \le -a_k$. In order to establish the first conclusion, observe that if $\sum_{\ell=1}^{j} b_\ell \le \sum_{\ell=1}^{j} b_\ell^*$ for $j = 1,2,\dots,k-1$ and $\sum_{\ell=1}^{k} b_\ell = \sum_{\ell=1}^{k} b_\ell^*$ then, setting $a_{k+1} = 0$, we can, by Abel's method of summation, write

$$\sum_{\ell=1}^{k} a_\ell b_\ell = \sum_{\ell=1}^{k} (a_\ell - a_{\ell+1}) \sum_{j=1}^{\ell} b_j \ge \sum_{\ell=1}^{k} (a_\ell - a_{\ell+1}) \sum_{j=1}^{\ell} b_j^*$$
$$= \sum_{\ell=1}^{k} a_\ell b_\ell^*.$$

Thus, since $P(\ell,k)$; $\ell = 1,2,\dots,k$, $\binom{k-1}{\ell-1}2^{-k+1}$; $\ell = 1,2,\dots,k$ and 1/2, 1/2, 0, ..., 0 are probability distributions on $\{1,2,\dots,k\}$ it suffices to show that $P(1,k) \le 1/2$ and

$$\sum_{\ell=1}^{j} \binom{k-1}{\ell-1}2^{-k+1} \le \sum_{\ell=1}^{j} P(\ell,k);\ j = 1,2,\dots,k-1. \tag{5}$$

The first inequality has already been established. (See (2).) In order to establish (5) we use induction on k and the pool adjacent violators algorithm (PAVA) for $\bar{\mu}_k$. For the case $k = 2$ it is easy to see that $P(1,2) = P(2,2) = 1/2$ independent of the values of w_1 and w_2. Assume (5) holds for k and let $L_k(\bar{X}_1,\bar{X}_2,\dots,\bar{X}_k; w_1,w_2,\dots,w_k)$ denote the number of level sets in the isotonic regression of $\bar{X}_1,\bar{X}_2,\dots,\bar{X}_k$ with weights $w_1,w_2,\dots,w_k$. Now, by the PAVA, the isotonic regression, $\bar{\mu}_{k+1}$, of $\bar{X}_1,\bar{X}_2,\dots,\bar{X}_{k+1}$ with weights $w_1,w_2,\dots,w_{k+1}$ may be formed by first constructing $\bar{\mu}_k$, the isotonic regression of $\bar{X}_1,\bar{X}_2,\dots,\bar{X}_k$ with weights $w_1,w_2,\dots,w_k$, and then combining $\bar{\mu}_k$ with $\bar{X}_{k+1}$ in the appropriate way. Thus, using the obvious notational abuses, either $L_{k+1} = L_k + 1$ or $L_{k+1} \le L_k$ and the former case is

characterized by $\bar{\mu}_{kk} < \bar{X}_{k+1}$. Thus, $L_{k+1} \leq L_k + I_{[\bar{X}_{k+1} > \bar{\mu}_k]}$ and for $j = 1,2,\ldots,k$,

$$P[L_{k+1} \geq j+1] \leq P[L_k + I_{[\bar{X}_{k+1} > \bar{\mu}_k]} \geq j+1] \tag{6}$$
$$= P[L_k = j \text{ and } \bar{X}_{k+1} > \bar{\mu}_{kk}] + P[L_k \geq j+1].$$

Using the notation in Barlow et al. (1972) and using the proof of their (3.23), we can express $P[L_k = j \text{ and } \bar{X}_{k+1} > \bar{\mu}_{kk}]$ as a sum of the form

$$\sum_{L_{j,k}} P[Av(B_1) < Av(B_2) < \ldots < Av(B_j) < \bar{X}_{k+1}] \prod_{i=1}^{j} P(1, C_{B_i}; w(B_i)).$$

Applying their inequality (3.20) to the first factor in each term yields the following bound:

$$\sum_{L_{j,k}} P[Av(B_j) < \bar{X}_{k+1}] P(j,j; W_{B_1}, W_{B_2}, \ldots, W_{B_j}) \prod_{i=1}^{j} P(1, C_{B_i}; w(B_i)).$$

Bounding the first factor in each term by 1/2 and again using (3.23) yields $P[L_k = j \text{ and } \bar{X}_{k+1} > \bar{\mu}_{kk}] \leq P(j,k)/2$. Thus, applying the induction hypothesis, we obtain

$$\begin{aligned} P[L_{k+1} \geq j+1] &\leq (P(j,k) + 2\,P[L_k \geq j+1])/2 \\ &= (P[L_k \geq j] + P[L_k \geq j+1])/2 \\ &\leq \sum_{\ell=j}^{k} \binom{k-1}{\ell-1} 2^{-k} + \sum_{\ell=j+1}^{k} \binom{k-1}{\ell-1} 2^{-k} \\ &= 2^{-k} \sum_{\ell=j+1}^{k+1} \left\{ \binom{k-1}{\ell-2} + \binom{k-1}{\ell-1} \right\} \\ &= 2^{-k} \sum_{\ell=j+1}^{k+1} \binom{k}{\ell-1}. \end{aligned}$$

This is the desired result.

It is conjectured on page 174 of Barlow et al. (1972)

that $\sum_{\ell=1}^{k} P(\ell,k)(-1)^{\ell} = 0$. In view of this conjecture it is interesting to note that both of the bounding distributions, $(1/2, 1/2, 0, 0, \ldots, 0)$ and $\binom{k-1}{\ell-1}(1/2)^{k-1}$; $\ell = 1,2,\ldots,k$, on $\{1,2,\ldots,k\}$ have this property.

The bounds given in (4) cannot, in general, be improved. Specifically, there exist sequences $w_n = (w_{n1},w_{n2},\ldots,w_{nk})$ and $w'_n = (w'_{n1},w'_{n2},\ldots,w'_{nk})$ with

$$\lim_{n\to\infty} P(\ell,k;\ w_n) = \binom{k-1}{\ell-1}2^{-k+1} \tag{7}$$

and

$$\lim_{n\to\infty} P(\ell,k;\ w'_n) = 2^{-1}I_{[\ell=1,2]} \tag{8}$$

for $\ell = 1,2,\ldots,k$. We first consider the shifted binomial distribution in (7). If we let

$$w(\varepsilon) = \begin{cases} (\varepsilon^{H},\varepsilon^{H-1},\ldots,\varepsilon,1,\varepsilon,\ldots,\varepsilon^{H}) & \text{for } k \text{ odd} \\ \\ (\varepsilon^{H-1},\varepsilon^{H-2},\ldots,\varepsilon,1,1,\varepsilon,\ldots,\varepsilon^{H-1}) & \text{for } k \text{ even} \end{cases} \tag{9}$$

where $H = [k/2]$, then (7) can be obtained by letting $\varepsilon\to 0$. The proof, however, is complicated and we offer a less involved proof by induction.

<u>Remark 2</u> Let $\mu_1 = \mu_2 = \ldots = \mu_k$. For each $k \geq 2$ and for each $\varepsilon > 0$ there exist positive weights $w_1,w_2,\ldots,w_k$ such that

$$\left|P(\ell,k;\ w_1,w_2,\ldots,w_k) - \binom{k-1}{\ell-1}(1/2)^{k-1}\right| < \varepsilon$$

for $\ell = 1,2,\ldots,k$.

<u>Proof</u>: Assume, without loss of generality, that the common value of $\mu_1,\mu_2,\ldots,\mu_k$ is zero. As before, the result is obvious

for $k = 2$. Assume that $k \geq 3$ and that $\varepsilon > 0$. By the induction hypothesis, there exist weights $w_1, w_2, \ldots, w_{k-1}$ such that

$$|P(\ell,k-1) - \binom{k-2}{\ell-1}(1/2)^{k-2}| < \varepsilon/4; \ \ell = 1,2,\ldots,k-1.$$

Let $w' = (w_1, w_2, \ldots, w_{k-1})$ and let $w = (w_1, w_2, \ldots, w_k)$ where w_k is to be specified. First consider $P(k,k;\ w)$. As noted in the proof of (3), $P(k,k;\ w) \leq (1/2)^{k-1}$. Write $P(k,k;\ w)$ as an iterated integral where the innermost integral is $\int_{x_{k-1}}^{\infty} (2\pi w_k^{-1})^{-1/2} \exp\{-x_k^2/(2w_k^{-1})\}dx_k$, which can be bounded below by $(1/2) - |x_{k-1}|(2\pi w_k^{-1})^{-1/2}$. Letting $w_k \to 0$ and applying the dominated convergence theorem, we obtain a lower bound which converges to $(1/2)P(k-1,k-1;\ w')$ as $w_k \to 0$. Thus, using our induction hypothesis, for sufficiently small w_k we have

$$(1/2)^{k-1} \geq P(k,k;\ w) \geq (1/2)\ P(k-1,k-1;\ w') - \varepsilon/2 \geq (1/2)^{k-1} - \varepsilon.$$

Now assume $\ell \leq k-1$ and write $P(\ell,k) = \sum_{L_{\ell k}}$ using (3.23) in Barlow et al. (1972). (For notational convenience we do not show the individual terms being summed). Partition the sum in (3.23) into terms involving B_ℓ such that $B_\ell = \{k\}$ and otherwise. Specifically, we write $P(\ell,k) = \sum_{L'_{\ell k}} + \sum_{L''_{\ell k}}$, where $L'_{\ell k}$ contains those elements in $L_{\ell k}$ with $B_\ell = \{k\}$ and $L''_{\ell k} = L_{\ell k} - L'_{\ell k}$. We consider $\sum_{L'_{\ell k}}$ first. The only factors, in the general term, which involve w_k are $P(1,1;\ w_k) = 1$ and $P(Av(B_1) < Av(B_2) < \ldots < Av(B_\ell))$. The argument given above for $P(k,k;\ w)$ applied to the first factor in each term of

$\sum_{L'_{\ell k}}$ shows that $\lim_{w_k \to 0} \sum_{L'_{\ell k}} = (1/2)P(\ell-1,k-1;\ w')$.

In considering $\lim_{w_k \to 0} \sum_{L''_{\ell k}}$, we first note that

$$P(Av(B_1) < Av(B_2) < \ldots < Av(B_\ell)) \to P(Av(B_1) < Av(B_2) < \ldots < Av(B_\ell - \{k\})).$$

If j is the smallest element of B_ℓ and if A denotes the event $[\min_{j \le \alpha \le k-1} Av(\{j,\ldots,\alpha\}) \ge Av(\{j,\ldots,k\})]$, then

$$P(1,C_{B_\ell};\ w(B_\ell)) = P(A \cap [\bar{X}_k > 0]) + P(A \cap [\bar{X}_k < 0]).$$

Furthermore,

$$\begin{aligned} P(A \cap [\bar{X}_k > 0]) &\le P[Av(\{j,\ldots,k-1\}) \ge Av(\{j,\ldots,k\}),\ \bar{X}_k > 0] \\ &\le P[Av(\{j,\ldots,k-1\}) \ge \bar{X}_k > 0] = P[\sqrt{w_k}\ Av(\{j,\ldots,k-1\}) \ge \sqrt{w_k}\ \bar{X}_k \ge \end{aligned}$$

but this probability converges to zero as $w_k \to 0$ since $\sqrt{w_k}\ \bar{X}_k$ has a standard normal distribution and $\sqrt{w_k}\ Av(\{j,\ldots,k-1\}) \to 0$. The other term, $P(A \cap [\bar{X}_k < 0])$, can be written as the sum of

$$P(A \cap [\bar{X}_k < 0,\ \min_{j \le \alpha \le k-1} Av(\{j,\ldots,\alpha\}) = Av(\{j,\ldots,k-1\})])$$

and

$$P(A \cap [\bar{X}_k < 0,\ \min_{j \le \alpha \le k-1} Av(\{j,\ldots,\alpha\}) < Av(\{j,\ldots,k-1\})]).$$

Since

$$\begin{aligned} &P(A^c \cap [\bar{X}_k < 0,\ \min_{j \le \alpha \le k-1} Av(\{j,\ldots,\alpha\}) = Av(\{j,\ldots,k-1\})]) \\ &\quad \le P[Av(\{j,\ldots,k\}) > Av(\{j,\ldots,k-1\}),\ \bar{X}_k < 0] \\ &\quad \le P[\sqrt{w_k}\ Av(\{j,\ldots,k-1\}) \le \sqrt{w_k}\ \bar{X}_k \le 0] \to 0 \text{ as } w_k \to 0, \end{aligned}$$

we see that

$$\begin{aligned} &P(A \cap [\bar{X}_k < 0,\ \min_{j \le \alpha \le k-1} Av(\{j,\ldots,\alpha\}) = Av(\{j,\ldots,k-1\})]) \to \\ &\quad (1/2)\ P[\min_{j \le \alpha \le k-1} Av(\{j,\ldots,\alpha\}) = Av(\{j,\ldots,k-1\})] = \\ &\qquad (1/2)\ P(1,C_{B_\ell - \{k\}};\ w(B_\ell - \{k\})). \end{aligned}$$

Note that $Av(\{j,\ldots,\alpha\})$ for $\alpha = j,\ldots,k-1$ does not depend on w_k and that $Av(B_\ell) = \sum_{\nu=j}^{k} w_\nu \bar{X}_\nu / \sum_{\nu=j}^{k} w_\nu \to Av(\{j,\ldots,k-1\})$ as $w_k \to 0$ and so $P(A \cap [\bar{X}_k < 0, \min_{j \leq \alpha \leq k-1} Av(\{j,\ldots,\alpha\}) = Av(\{j,\ldots,k-1\})]) \to 0$ as $w_k \to 0$. Hence, $\lim_{w_k \to 0} \sum_{L''_{\ell k}} = (1/2) P(\ell; k-1; w'$
Thus, $\lim_{w_k \to 0} P(\ell,k;\ w) = (P(\ell-1,k-1;\ w') + P(\ell,k-1;\ w'))/2$. By our induction hypothesis, this limit is sufficiently close to $\binom{k-2}{\ell-2}(1/2)^{k-1} + \binom{k-2}{\ell-1}(1/2)^{k-1} = \binom{k-1}{\ell-1}(1/2)^{k-1}$.

<u>Remark 3</u> Let $\mu_1 = \mu_2 = \ldots = \mu_k$. If $w'_n = (w'_{n1}, w'_{n2}, \ldots, w'_{nk}) \to (a,0,0,\ldots,b)$ with a and b positive, then $P(\ell,k;\ w'_n) \to 2^{-1}\, I_{[\ell=1,2]}$ for $\ell = 1,2,\ldots,k$.

<u>Proof</u>: Without loss of generality, we assume that $\bar{X}_i = Z_i/\sqrt{w'_{ni}}$ for $i = 1,2,\ldots,k$ with Z_i fixed, independent, identically distributed, standard normal variables. Clearly, the random vector $(\sqrt{w'_{ni}}\, Z_1, \sqrt{w'_{n2}}\, Z_2, \ldots, \sqrt{w'_{nk}}\, Z_k, w'_{n1}, \ldots, w'_{nk})$ converges weakly to $(\sqrt{a}\, Z_1, 0, 0, \ldots, 0, \sqrt{b}\, Z_k, a, 0, \ldots, 0, b)$ and so $Av(\{1,2,\ldots,k\}) - \max_{1 \leq i \leq k-1} Av(\{1,2,\ldots,i\})$ converges weakly to $b[(Z_k/\sqrt{b}) - (Z_1/\sqrt{a})]/(a+b)$. Hence,
$P(1,k;\ w'_n) = P[Av(\{1,2,\ldots,k\}) \geq \max_{1 \leq i \leq k-1} Av(\{1,2,\ldots,i\})] \to 1/$
Next consider $P(\ell,k)$ for $\ell \geq 3$. Write $P(\ell,k)$ using (3.23) in Barlow et al. (1972) and bound the general term by
$P[Av(B_1) < A(B_2) < \ldots < Av(B_\ell)] \leq P[Av(B_1) < Av(B_2 \cup B_3 \cup \ldots \cup B_{\ell-1}) < Av(B_\ell)]$. (This inequality follows from the Cauchy Mean Value Property of averages). Now this bound can be written as
$P[\sigma_1 Z_1 < \sigma_2 Z_2 < \sigma_3 Z_3] = P[(\sigma_1/\sigma_2) Z_1 < Z_2 < (\sigma_3/\sigma_2) Z_3]$ where $\sigma_1/\sigma_2 \to 0$ and $\sigma_3/\sigma_2 \to 0$. It follows that $\lim_{n\to\infty} P(\ell,k;\ w'_n) = 0$ for $\ell \geq 3$.

The number, L_k, of level sets in $\bar{\mu}_k$ has been suggested as a test statistic for the testing problem discussed in the introduction. Theorem 1 provides stochastic bounds on the distribution of L_k.

3. APPLICATIONS.

Example 1 Consider the problem, discussed in the introduction, of testing the homogeneity of a collection of normal means against the trend $\mu_1 \leq \mu_2 \leq \cdots \leq \mu_k$. For $c > 0$, Theorem 1 gives

$$(10)\quad P[\chi_1^2 \geq c]/2 \leq P[T_{01} \geq c] \leq \sum_{\ell=2}^{k} \binom{k-1}{\ell-1} 2^{-k+1} P[\chi_{\ell-1}^2 \geq c].$$

These bounds, which may be quite different, are tabled in Table 1 for k = 5,8,12 and various values of c. The values of $P[T_{01} \geq c]$ under the assumption of equal weights are also included. The equal weights probabilities seem to be somewhere around the average of the two bounds and may provide a reasonable approximation provided the weights are not one of the extreme types suggested by the "sharpness" analysis of Section 2. For example, if k = 5, if we observe a value of T_{01} near 5 and if we use the tail probabilities associated with equal weights, then we would report a p-value near .05. The actual p-value could be as large as .1 but not smaller than .013. Using the upper bound, one can show that a test whose significance level is no larger than .05 regardless of the weights is obtained by rejecting for values of T_{01} larger than 6.5. In this example, it should be noted that for k = 8 and especially k = 12 there is

considerable difference between the upper and lower bounds.

The "sharpness" analysis provided by Remarks 2 and 3 suggests that $P[T_{01} \geq c]$ is close to the upper bound when the weights associated with the extreme populations (close to k or 1) are small. On the other hand, it should be close to the lower bound when the weights on the extreme populations are relatively large. Since the errors in the estimators of μ_i for extreme values of i are larger than those for middle values of i (cf. Cryer, Robertson, Wright and Casedy (1972)), one might actually sample more items from these extreme populations creating the latter type of situation.

If the population variances are unknown but of the form $\sigma_i^2 = \sigma^2 a_i$ with a_i known, then the likelihood ratio statistic, $S_{01} = 1 - \lambda^{2/N}$ for testing homogeneity against trend satisfies

$$P[S_{01} \geq c] = \sum_{\ell=1}^{k} P(\ell,k)P[B_{(\ell-1)/2,\ (N-\ell)/2} \geq c]$$

where $N = \sum_{i=1}^{k} n_i$, $B_{a,b}$ is a beta variable with parameters a and b, $B_{0b} \equiv 0$, and the weights for determining $P(\ell,k)$ are $w_i = n_i/a_i$; $i = 1,2,\ldots,k$ (cf. Barlow et al. (1972)). Now, for $\ell > 1$, $(U_1 + \ldots + U_{\ell-1})/(U_1 + \ldots + U_{N-1})$ has a $B_{(\ell-1)/2,\ (N-\ell)/2}$ distribution provided $U_1, U_2, \ldots, U_{N-1}$ are i.i.d gamma variables with $\alpha = 1/2$ and common β. Thus $P[B_{(\ell-1)/2,(N-\ell)/2} \geq c]$ is increasing in ℓ and Theorem 1 applies. If $c > 0$ and H_0 is true, we obtain

$$P[B_{1/2,\ (N-2)/2} \geq c]/2 \leq P[S_{01} \geq c] \leq \sum_{\ell=2}^{k} \binom{k-1}{\ell-1} 2^{-k+1} P[B_{(\ell-1)/2,(N-\ell)/2} \geq c]. \tag{11}$$

<u>Example 2</u> Again suppose $\mu_1, \mu_2, \ldots, \mu_k$ are the means of k normal populations. Robertson and Wegman (1978) considered likelihood ratio tests of the null hypothesis H_1: $\mu_1 \leq \mu_2 \leq \ldots \leq \mu_k$. The configuration H_0: $\mu_1 = \mu_2 = \ldots = \mu_k$ is least favorable within H_1 so that conservative tests are obtained using critical regions computed under H_0. If the variances are known then the distribution, under H_0, of $T_{12} = -2 \ln \lambda = \sum_{i=1}^{k} w_i (\bar{\mu}_{ki} - \bar{X}_i)^2$ is again a chi-bar-squared distribution ($w_i = n_i/\sigma_i^2$). Specifically, under H_0,

$$P[T_{12} \geq c] = \sum_{\ell=1}^{k} P(\ell,k) \; P[\chi^2_{k-\ell} \geq c].$$

Theorem 1 gives the following bounds

$$(12) \quad \sum_{\ell=1}^{k} \binom{k-1}{\ell-1} 2^{-k+1} P[\chi^2_{k-\ell} \geq c] \leq P[T_{12} \geq c] \leq (P[\chi^2_{k-2} \geq c] + P[\chi^2_{k-1} \geq c]).$$

These bounds, together with the probabilities computed for equal weights, are given in Table 2 for k = 5, 8 and 12. Unlike the last example, for even as many twelve populations the tail probabilities for the equal weights case are not too different from those of the upper bound. Thus, computing probabilities under the equal weights case is recommended provided the extreme weights are not relatively large or small. The upper bound in (12) can be used to determine conservative tests of H_1.

If the variances are unknown but of the form $\sigma_i^2 = \sigma^2 a_i$ with a_i known then the likelihood ratio statistic, $S_{12} = 1 - \lambda^{2/N}$, for testing H_1 against all alternatives has the following distribution, under H_0,

$$P[S_{12} \geq c] = \sum_{\ell=1}^{k} P(\ell,k) \; P[B_{(k-\ell)/2, \; (N-k)/2} \geq c].$$

Again let $U_1, U_2, \ldots, U_{N-\ell}$ be independent gamma variables with $\alpha = 1/2$ and common β. Let $X = \sum_{i=1}^{k-\ell-1} U_i$, $Y = U_{k-\ell}$ and $Z = \sum_{i=k-\ell+1}^{N-\ell} U_i$. Since $B_{(k-\ell-1)/2,\ (N-k)/2} \sim X/(X+Z)$, $B_{(k-\ell)/2,\ (N-k)/2} \sim (X+Y)/(X+Y+Z)$ and $(x+y)/(x+y+z) \geq x/(x+z)$ for all $x,y,z \geq 0$, Theorem 1 also applies to this test.

Robertson and Wegman (1978) also consider likelihood ratio tests when the sampled populations belong to an exponential family and when an order restriction is imposed on the parameters indexing the populations by either the null or the alternative hypothesis. Chi-bar-squared distributions arise as limit distributions of the test statistics and Theorem 1 applies. Robertson (1978) considered likelihood ratio statistics for tests involving order restrictions on a collection of multinomial parameters. Again, chi-bar-squared distributions arise as limiting distributions and Theorem 1 applies.

Example 3 Suppose we have two multinomial populations with parameter sets $p = (p_1, p_2, \ldots, p_k)$ and $q = (q_1, q_1, \ldots, q_k)$. Suppose we have independent samples from each of these populations and we wish to test H_0: $p = q$ against the alternative $H_1 - H_0$ where

$$H_1:\ p_1 \geq q_1,\ p_1 + p_2 \geq q_1 + q_2, \ldots, p_1 + p_2 + \ldots + p_{k-1} \geq q_1 + q_2 + \ldots + q_{k-1}$$

Under H_0, the likelihood ratio statistic, $-2 \ln \lambda$, has an asymptotic distribution with tail probabilities of the form $\sum_{\ell=1}^{k} P(\ell,k)\, P[\chi^2_{k-\ell} \geq c]$ where the $P(\ell,k)$ are determined by the weights $p_1, p_2, \ldots, p_k$. Of course, the p vector is unknown under H_0 so that in order to apply this result a least favorable

configuration is required. Theorem 1 provides this configuration. These results together with several other testing problems are considered in Robertson and Wright (1980).

REFERENCES

Barlow, R. E., Bartholomew, D. J., Bremner, J. M. and Brunk, H. D. (1972). Statistical Inference Under Order Restrictions. Wiley, New York.

Bartholomew, D. J. (1959). A test for homogeneity for ordered alternatives I and II. Biometrika 46 36-48, 328-335.

Cryer, J. D., Robertson, Tim, Wright, F. T. and Casady, R. J. (1972). Monotone median regression. Ann. Math. Statist. 43 1459-1469.

Grove, D. M. (1980). A test of independence against a class of ordered alternatives in a 2 × C contingency table. J. Amer. Statist. Assoc. 75 454-459.

Robertson, Tim and Wegman, Edward J. (1978). Likelihood ratio tests for order restrictions in an exponential families. Ann. Statist. 6 485-505.

Robertson, Tim (1978). Testing for an against an order restriction on multinomial parameters. J. Amer. Statist. Assoc. 73 197-202.

Robertson, Tim and Wright, F. T. (1980) Likelihood ratio tests for and against a stochastic ordering between multinomial populations. (Submitted).

Table 1. The lower bound (LB) and upper bound (UB) for $P[T_{01} \geq c]$ and the value of $P[T_{01} \geq c]$ for equal weights (EW).

	k = 5			k = 8			k = 12		
	c=3.6	c=5.0	c=8.3	c=4.5	c=6.1	c=9.7	c=5.3	c=7.0	c=10.7
LB	.029	.013	.002	.017	.007	.001	.011	.004	.001
EW	.102	.051	.010	.102	.050	.010	.100	.049	.010
UB	.182	.098	.022	.289	.166	.043	.441	.284	.094

Table 2. The lower bound (LB) and upper bound (UB) for $P[T_{12} \geq c]$ and the value of $P[T_{12} \geq c]$ for equal weights (EW).

	k = 5			k = 8			k = 12		
	c=6.1	c=7.6	c=11.3	c=9.9	c=11.9	c=16.1	c=14.8	c=17.1	c=22.0
LB	.060	.030	.006	.040	.018	.003	.023	.010	.002
EW	.098	.051	.010	.100	.049	.010	.100	.050	.010
UB	.150	.081	.017	.161	.084	.019	.166	.089	.020

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: 14 1
2. GOVT ACCESSION NO.: AD-A087117
3. RECIPIENT'S CATALOG NUMBER: 9
4. TITLE (and Subtitle): 6 BOUNDS ON MIXTURES OF DISTRIBUTIONS ARISING IN ORDER RESTRICTED INFERENCE; (Distributional bounds).
5. TYPE OF REPORT & PERIOD COVERED: TECHNICAL REPORT,
6. PERFORMING ORG. REPORT NUMBER:
7. AUTHOR(s): 10 Tim Robertson [redacted] F. T. Wright
8. CONTRACT OR GRANT NUMBER(s): 15 ~~N00014-80-C-0321~~ N00014-80-C-0322
9. PERFORMING ORGANIZATION NAME AND ADDRESS: Dept. of Statistics, University of Iowa, Iowa City, Iowa 52242; Dept. of Mathematics, Univ. of Missouri-Rolla, Rolla, MO. 65401
10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS:
11. CONTROLLING OFFICE NAME AND ADDRESS: Office of Naval Research, Statistics and Probability Program Code 436, Arlington, Virginia 22217
12. REPORT DATE: July 10, 1980
13. NUMBER OF PAGES: 20
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office): 11 10 Jul 80
15. SECURITY CLASS. (of this report): UNCLASSIFIED
15a. DECLASSIFICATION/DOWNGRADING SCHEDULE:
16. DISTRIBUTION STATEMENT (of this Report): APPROVED FOR PUBLIC RELEASE: DISTRIBUTION UNLIMITED. 12 22
17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report):
18. SUPPLEMENTARY NOTES:
19. KEY WORDS (Continue on reverse side if necessary and identify by block number): Order restricted inference, tests for and against a trend, chi-bar-squared distribution, E-bar-squared distribution, tail probability bounds, least favorable configurations.
20. ABSTRACT (Continue on reverse side if necessary and identify by block number): PLEASE SEE REVERSE SIDE

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

S/N 0102-LF-014-6601

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

345 2??

ABSTRACT

In testing hypotheses involving order restrictions on a collection of parameters, distributions arise which are mixtures of standard distributions. Since tractable expressions for the mixing proportions do not, in general, exist even for parameter collections of moderate size, the implementation of these tests may be difficult. It is not necessary to bound the individual mixing proportions in order to obtain a bound for the tail probabilities of the "mixed" distributions. Bounds for such tail probabilities are derived and their application to several order restricted hypothesis testing situations is discussed. These results are also applied to obtain the least favorable configuration in a two sample likelihood ratio test of the equality of two multinomial populations versus a stochastic ordering alternative.